



Hierarchical log Gaussian Cox process for regeneration in uneven-aged forests

Downloaded from: <https://research.chalmers.se>, 2023-05-04 22:15 UTC

Citation for the original published paper (version of record):

Kuronen, M., Särkkä, A., Vihola, M. et al (2022). Hierarchical log Gaussian Cox process for regeneration in uneven-aged forests. *Environmental and Ecological Statistics*, 29(1): 185-205.
<http://dx.doi.org/10.1007/s10651-021-00514-3>

N.B. When citing this work, cite the original published paper.



Hierarchical log Gaussian Cox process for regeneration in uneven-aged forests

Mikko Kuronen¹ · Aila Särkkä² · Matti Vihola³ · Mari Myllymäki¹

Received: 12 April 2021 / Revised: 6 July 2021 / Accepted: 26 July 2021

© The Author(s) 2021

Abstract

We propose a hierarchical log Gaussian Cox process (LGCP) for point patterns, where a set of points \mathbf{x} affects another set of points \mathbf{y} but not vice versa. We use the model to investigate the effect of large trees on the locations of seedlings. In the model, every point in \mathbf{x} has a parametric influence kernel or signal, which together form an influence field. Conditionally on the parameters, the influence field acts as a spatial covariate in the intensity of the model, and the intensity itself is a non-linear function of the parameters. Points outside the observation window may affect the influence field inside the window. We propose an edge correction to account for this missing data. The parameters of the model are estimated in a Bayesian framework using Markov chain Monte Carlo where a Laplace approximation is used for the Gaussian field of the LGCP model. The proposed model is used to analyze the effect of large trees on the success of regeneration in uneven-aged forest stands in Finland.

Keywords Bayesian inference · Competition kernel · Laplace approximation · MCMC · Spatial random effects · Tree regeneration

Mathematics Subject Classification 62M30 · 62F15 · 60G55

Handling Editor: Luiz Duczmal.

Mikko Kuronen, Mari Myllymäki and Matti Vihola were financially supported by the Academy of Finland (Project Numbers 306875, 327211, 295100 and 315619) and Aila Särkkä by the Swedish Research Council (VR 2018-03986).

✉ Mikko Kuronen
mikko.kuronen@luke.fi

¹ Natural Resources Institute Finland (Luke), Helsinki, Finland

² Mathematical Sciences, Chalmers University of Technology and the University of Gothenburg, Gothenburg, Sweden

³ Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä, Finland

1 Introduction

Hierarchical relationships or interactions, where a plant species affects the locations or intensity of another species but not vice versa, often occur in ecological communities (e.g. Dieckmann et al. 2000). An example of such a hierarchical relationship is that proximity of large trees affects the intensity of seedling either positively, e.g. by protecting against wind, or negatively by giving too much shade. Mathematically, we can describe such plant communities by two point processes, Y and X , where one (X) is affecting the other (Y) but not vice versa.

The hierarchical interaction assumption affects the inference for Y and X greatly since X can be modeled independently of Y and Y is modeled conditionally on X . A realization of the point process X acts then as a source of heterogeneity in the distribution of Y . Högmader and Särkkä (1999) modeled interaction between two territorial ant species using Gibbs point processes under such an assumption. A similar hierarchical Gibbs point process approach was used in Grabarnik and Särkkä (2009) and Genet et al. (2014). Furthermore, Illian et al. (2009) modeled the spatial pattern of resprouter species (Y) given the locations of seeders (X) in a hierarchical set-up having an inhomogeneous Poisson process as a model for the resprouters.

Here, we model the intensity of new seedlings in a spruce-dominated uneven-aged (boreal) forest given the locations and diameters at breast height (dbh) of large trees. Thus, our X process of large trees is a marked point process, where the mark of a tree is the dbh. The data consist of 14 sample plots from an experiment of continuous cover forestry involving single-tree selection in four nearby areas in Southern Finland (Fig. 1). The system relies on the natural emergence of new seedlings, and a continuous recruitment is necessary for long-term sustainability in a wide sense (e.g. Eerikäinen et al. 2014; Kuusinen et al. 2019). While a sufficient number of seedlings is necessary for the success of regeneration, our focus here is on the spatial distribution of the seedlings within the plots, and the effect of large trees on it.

Like in the resprouter and seeder case above, an inhomogeneous Poisson process would be a reasonable model since the effect of large trees could be added in the model as an explanatory variable. However, already visual inspection of the patterns of seedlings y indicates that the patterns tend to be rather clustered, beyond the clustering that may be explained by the patterns of large trees x . Due to such unexplained clustering, a log Gaussian Cox process (LGCP) (Møller et al. 1998) is a more appropriate model for the conditional point process of seedlings given large trees.

To model the effect of the large trees X , we assume that each tree $x \in X$ emits a signal or impulse that describes the effect of the tree on its neighborhood. We assume that this effect decreases with the distance from the tree x . In general, the size of the effect as well as the range of the effect could depend on the size or other properties of the tree, e.g. its dbh. Because we do not have precise a priori information on the size and range of the effects, we use parametric signals similar to the ones found in the literature (Adler 1996; Pommerening et al. 2011; Häbel et al. 2019; Pommerening and Grabarnik 2019). The individual signals are then superimposed to form an influence field, which describes the overall influence of the points of X on any location s in the observation window W . These kinds of models have been used to model, for example, effect of neighboring individuals on the growth of a subject tree, survival of seedlings

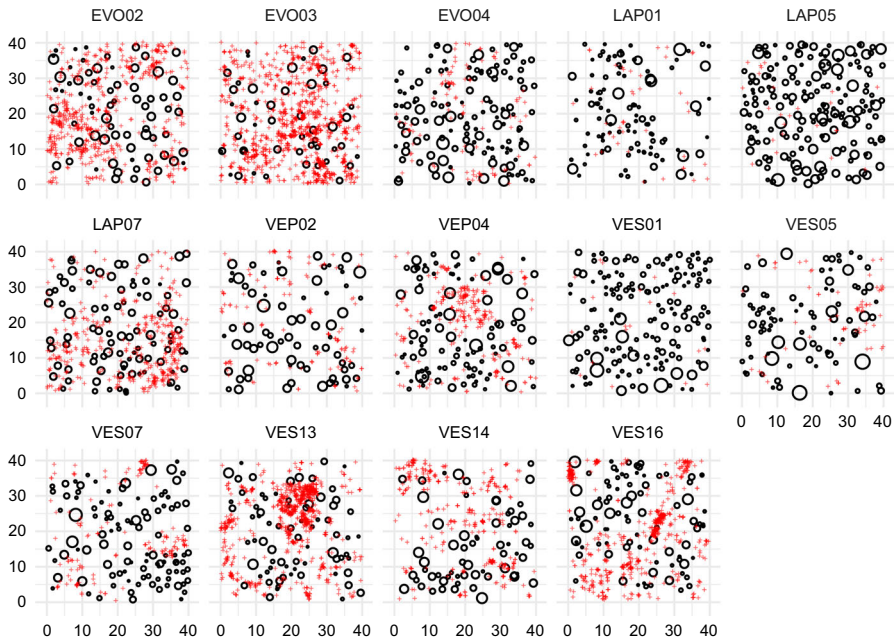


Fig. 1 Trees with dbh at least 7 cm (open circles with radii relative to the dbh of the tree) and new seedlings (red crosses) in areas of size 40 m \times 40 m. The headings give abbreviations for the plot locations and numbers

and ground vegetation in different contexts (e.g. Wu et al. 1985; Miina and Pukkala 2002; Pommerening et al. 2011; Häbel et al. 2019; Kuuluvainen and Pukkala 1989; Köhlmann-Berenzon et al. 2005).

Our idea here is to include the superimposed individual signals in the log intensity function of the LGCP model. Using parametric models for the signals, the intensity of the LGCP is a non-linear function of the model parameters. According to Pommerening and Sánchez Meador (2018) the signals are aggregated additively or multiplicatively and there is no evidence to prefer either of these ways. We follow Pommerening et al. (2011) and Illian et al. (2009) and aggregate the signals additively.

Our Bayesian inference algorithm is based on Markov chain Monte Carlo (MCMC) sampling for parameters, and a Laplace approximation is used for the latent random field of the LGCP to avoid high-dimensional MCMC sampling. Laplace approximations are widely used for inference of latent Gaussian fields, for instance within the popular INLA method (Rue et al. 2009). However, in contrast to INLA, MCMC is more robust, and can cope with multimodal parameter posteriors.

The large tree process typically extends beyond the borders of the sample plot. However, we have observed the process in the same observation window as the seedlings. Thus, the influence field computed only from the observed trees is weaker near the borders than the field computed from the fully observed large tree process would be. In order to account for the unobserved trees outside the observation window, we compute the influence field using an edge correction method similar to that suggested

in Kühnmann-Berenzon et al. (2005): the unobserved trees are imputed based on the assumption that the locations of large trees are distributed according to a Poisson process. This rather simple edge correction method can be efficiently implemented within the Bayesian inference, in contrast to alternatives where the locations (and sizes) of unobserved large trees would be included in the Bayesian inference as unknowns and simulated within the MCMC approach.

The rest of the paper is organized as follows. In Sect. 2, we give some examples of influence kernels and introduce the conditional LGCP model. The Bayesian estimation approach, including the edge correction, is described in Sect. 3. Section 4 presents the results of a simulation experiment that was conducted to explore the performance of the proposed estimation and edge correction methods. Finally, the forestry data are described in further detail and studied in Sect. 5. Section 6 is for discussion.

2 Conditional log Gaussian Cox process model

Let us have a bivariate point process in \mathbb{R}^2 consisting of an unmarked point process Y and an unmarked or a marked point process X . Let us further assume that we have observed a realization of process Y , namely $\mathbf{y} = \{y_i\}$, in a bounded window $W \subset \mathbb{R}^2$. Our primary interest is in the spatial pattern \mathbf{y} which is affected by a realization \mathbf{x} of the spatial point process X . The spatial pattern \mathbf{x} can consist only of the point locations x_j or of the point locations and marks, $[x_j, m_j]$, if some characteristics (marks) m_j of the points x_j are available. In our forestry application, \mathbf{y} consists of the locations of seedlings, while \mathbf{x} is the pattern of locations and dbh's of large trees.

In our approach, the effect of \mathbf{x} on \mathbf{y} is modeled using the influence kernels around the points of \mathbf{x} that are explained in Sect. 2.1. To account for the clustering in the pattern \mathbf{y} not explained by \mathbf{x} , the LGCP model is proposed and defined in Sect. 2.2. Replicated point patterns are discussed in Sect. 2.3.

2.1 Influence kernels and influence field

We assume that each point $[x_j, m_j]$ of the process X introduces an influence kernel around its location. We focus on isotropic influence kernels of the form $c(h; m_j, \theta_I)$, where $h = \|s - x_j\|$ is the distance between the location s of interest and x_j . Many kernels have been suggested in the literature for different applications (e.g. Adler 1996; Illian et al. 2008; Pommerening et al. 2011; Pommerening and Maleki 2014; Schneider et al. 2006). We used a mark independent Gaussian kernel

$$c(h; \theta) = \exp\left(-(h/\theta)^2\right), \quad (1)$$

where $\theta > 0$ is an unknown influence range parameter. Here the influence of a point gradually decreases with the distance from the point.

A mark dependent generalization of (1) is given by

$$c(h, m; \theta_I) = m^\alpha \exp\left(-\left(\frac{h}{\theta m^\delta}\right)^2\right) \quad (2)$$

with $\theta_I = (\theta, \delta, \alpha)$, where $\theta > 0$, $\delta > 0$, and $\alpha \geq 0$. If $\alpha = 0$, the mark affects only the range of influence and if $\alpha > 0$, it affects both the range and the strength (e.g. Pommerening et al. 2011).

The influence field of the process X can then be defined as a superposition of the individual influence kernels,

$$C(s; \theta_I, X) = \sum_{[x_j, m_j] \in X} c(\|s - x_j\|, m_j; \theta_I).$$

2.2 Conditional model

Since y is affected by x , we introduce a conditional point process model for y given $X = x$, where the intensity of Y is affected by the influence field of x . This conditional model is a LGCP with the intensity

$$\Lambda(s; \beta, \theta_I, x, Z) = \exp(\beta_0 + \beta_1 C(s; \theta_I, x) + Z(s)), \quad (3)$$

where $C(s; \theta_I, x)$ is a parametric influence field, $\beta = (\beta_0, \beta_1)$ and the unknown coefficients $\beta_0 \in \mathbb{R}$ and $\beta_1 \in \mathbb{R}$ are the intercept and the strength of the influence field, respectively. If $\beta_1 < 0$, x affects the intensity of Y negatively and the influence field $C(s; \theta_I, x)$ can be interpreted as a thinning of the LGCP process with intensity $\Lambda(s) = \exp(\beta_0 + Z(s))$. If, however, $\beta_1 > 0$, x has a positive effect on the intensity of Y and there are more points of Y in areas with a high value of $C(s; \theta_I, x)$. Furthermore, $Z := \{Z(s) : s \in \mathbb{R}^2\}$ is a zero-mean stationary Gaussian random field with a covariance function $C_Z(r; \theta_Z)$ and independent of the influence field. In our application below, we use the Matérn covariance function

$$C_Z(r; \theta_Z, \nu) = \sigma_Z^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} \frac{r}{\rho_Z} \right)^\nu K_\nu \left(\sqrt{2\nu} \frac{r}{\rho_Z} \right), \quad r > 0, \quad (4)$$

with the smoothness parameter $\nu = 2$ and $\theta_Z = (\sigma_Z^2, \rho_Z)$, where σ_Z^2 and ρ_Z are the variance and range parameters, respectively, and K_ν is the modified Bessel function of the second kind (e.g. Cressie 1993; Chilés and Delfiner 1999; Banerjee et al. 2004). The choice $\nu = 2$ was made since we expect that the unobserved environmental conditions that affect the clustering of y in our application vary rather smoothly and since it is computationally convenient (Lindgren et al. 2011).

2.3 Replicates

Assume that we have several independent replicated point patterns $y_k, k = 1, \dots, N$, from the conditional distribution of the point process Y given $X = x_k, k = 1, \dots, N$. Conditionally on $X = x_k$, the model for y_k is a LGCP with the intensity $\Lambda(s; \beta_0, \beta_1, \theta_I, x_k, Z_k)$ in (3), where $Z_k, k = 1, \dots, N$, are independent replicates of the Gaussian random field with parameters θ_Z . For our data, it is not reasonable to assume that all replicates have the same β_0 , which controls the number of points

of Y , and we let each pattern y_k have its own intercept parameter β_0 , i.e. β_{0k} for y_k , $k = 1, \dots, N$. Consequently, in our application below, the pattern y_k is assumed to be a realization of the LGCP model with the intensity $\Lambda(s; \beta_{0k}, \beta_1, \theta_I, x_k, Z_k)$.

3 Inference

The likelihood of the conditional LGCP model for a point pattern y with n points observed in W is

$$p(y; \beta, \theta_I, \theta_Z, x) = \mathbf{E}_{\theta_Z} \prod_{i=1}^n \Lambda(y_i; \beta, \theta_I, x, Z) \exp \left(- \int_W \Lambda(u; \beta, \theta_I, x, Z) du \right), \quad (5)$$

where $\beta, \theta_I, \theta_Z$ are the model parameters, Z denotes the Gaussian random field and the expectation is over Z given θ_Z . As we use Bayesian inference we need to be able to evaluate the likelihood (5) efficiently. Below, we describe the approximations needed: discretization of the observation window (Sect. 3.1), an edge-corrected influence field (Sect. 3.2), and approximations related to the Gaussian field (Sect. 3.3), which include approximating the field by a Gaussian Markov random field and using the Laplace approximation to evaluate the likelihood. Finally, the approximated likelihood based on replicates is given in Sect. 3.4 and the MCMC algorithm is described in Sect. 3.5.

3.1 Discretization

To be able to make inference on LGCP models, the observation window W of the point pattern y is discretized using a regular grid in a similar manner as in Rue et al. (2009) and Møller et al. (1998). Namely, the observation window W is divided into G disjoint cells $\{w_g\}$ with center locations ξ_g and area A . Furthermore, we let n_g^y denote the number of observations y within w_g in W and $\mathbf{n}^y = (n_1^y, \dots, n_G^y)$. A piecewise constant approximation is used for the Gaussian field Z and the competition field C and the locations of y are replaced by the counts n_g^y . The approximate likelihood for \mathbf{n}^y is

$$p(\mathbf{n}^y; \beta, \theta_I, \theta_Z, x) = \mathbf{E}_{\theta_Z} p(\mathbf{n}^y; \beta, \theta_I, x, Z^D), \quad (6)$$

where

$$p(\mathbf{n}^y; \beta, \theta_I, x, Z^D) = \prod_{g=1}^G \text{Pois}(n_g^y; \Lambda_g),$$

$\Lambda_g = A \exp(\beta_0 + \beta_1 C^D(\xi_g; \theta_I, x) + Z^D(\xi_g; \theta_Z))$, and C^D and Z^D are the piecewise constant approximations of C and Z .

3.2 Edge correction

The large tree process X is only partially observed, and generating the influence field only based on the observed large trees would result in too weak influence near the borders. Therefore, we propose an imputation type approach, similar to the one proposed by Kühnmann-Berenzon et al. (2005), to correct for the unobserved points of X . Specifically, we propose to replace the influence generated by the unobserved trees with the expected influence generated assuming that the whole process X is an independently marked homogeneous Poisson process. In the unmarked case, X is assumed to be a homogeneous Poisson process. In general, the point pattern outside the window would depend on the pattern inside the window, but this is not the case for the Poisson process.

Let λ and F be the intensity and mark distribution of X , and X_{W^c} the restriction of X to W^c , the complement of W . Using the Campbell theorem (e.g. Chiu et al. 2013) we can write

$$\begin{aligned} \mathbf{EC}(s; \theta_I, X_{W^c}) &= \mathbf{E} \sum_{[x_j, m_j] \in X_{W^c}} c(\|s - x_j\|, m_j; \theta_I) \\ &= \int_{R_+} \int_{R^2} c(s - x, m; \theta_I) \mathbf{1}_{W^c}(x) \lambda \, dx \, dF(m), \end{aligned}$$

where $\mathbf{1}_{W^c}$ is the indicator function of the set W^c , i.e. $\mathbf{1}_{W^c}(x) = 1$ if $x \in W^c$, and 0 otherwise. By changing the order of the integrals we find that

$$\begin{aligned} \mathbf{EC}(s; \theta_I, X_{W^c}) &= \int_{R^2} f(s - y) \mathbf{1}_{W^c}(y) \lambda \, dy \\ &= \int_{R^2} f(s - y) \lambda \, dx - \int_{R^2} f(s - x) \mathbf{1}_W(x) \lambda \, dx, \end{aligned}$$

where $f(x) = \int_{R_+} c(\|x\|, m; \theta_I) \, dF(m)$. By changing to polar coordinates and with a slight abuse of notation

$$\int_{R^2} f(s - x) \lambda \, dx = \lambda 2\pi \int_0^\infty r f(r) \, dr,$$

which can be computed using numerical integration. Since we are only interested in locations $s \in W$, we can replace the function f with $f \mathbf{1}_{W^S}$, the restriction of f to the set $W^S = \{s - x : s \in W, x \in W\}$, and

$$\int_{R^2} f(s - x) \mathbf{1}_W(x) \, dx = \int_{R^2} (f \mathbf{1}_{W^S})(s - x) \mathbf{1}_W(x) \lambda \, dx = (f \mathbf{1}_{W^S} * \mathbf{1}_W)(s).$$

The discrete convolution of the piecewise constant approximations of $f \mathbf{1}_{W^S}$, and $\mathbf{1}_W$ can be efficiently computed using discrete Fourier transforms (Oppenheim et al. 1999; Frigo and Johnson 2005). For F , we use the empirical distribution of marks in the sample plot under study.

The edge-corrected influence field value at any location $s \in W$ is then obtained as the sum of the influence field calculated from the observed \mathbf{x}_W , $C(s; \boldsymbol{\theta}_I, \mathbf{x}_W)$, and the expected influence load of the unobserved X_{W^c} . In general, we use the numerical approximation explained above, but for the special case of the Gaussian influence kernel (1) and a rectangular observation window, it is easy to compute the edge correction by hand.

3.3 Approximations related to the Gaussian field

We use Laplace approximation (Tierney and Kadane 1986; Rue et al. 2009) to approximate the likelihood (6) and obtain

$$\mathbf{E}_{\boldsymbol{\theta}_Z} p(\mathbf{n}^y; \boldsymbol{\beta}, \boldsymbol{\theta}_I, \mathbf{x}, Z^D) \approx \sqrt{\frac{(2\pi)^d}{\det(-\mathbf{H}(\hat{\mathbf{z}}))}} p(\mathbf{n}^y; \boldsymbol{\beta}, \boldsymbol{\theta}_I, \mathbf{x}, \hat{\mathbf{z}}) p(\hat{\mathbf{z}}; \boldsymbol{\theta}_Z), \quad (7)$$

where \mathbf{H} and $\hat{\mathbf{z}}$ are the Hessian and maximizer of $\log p(\mathbf{n}^y; \boldsymbol{\beta}, \boldsymbol{\theta}_I, \mathbf{x}, \mathbf{z}) p(\mathbf{z}; \boldsymbol{\theta}_Z)$, respectively, and $p(\mathbf{z}; \boldsymbol{\theta}_Z)$ is the probability density of the vector \mathbf{Z}^D which contains the values of Z^D at grid cells.

Since the Gaussian random field Z is assumed to have mean zero and the Matérn covariance function (4) with $\nu = 2$, we can utilize the explicit link between Gaussian fields and Markov random fields (Lindgren et al. 2011), which tells us that the distribution of \mathbf{Z}^D should be approximated with a Gaussian distribution with a precision matrix given by Lindgren et al. (2011).

3.4 Replicates

Since the point patterns are assumed to be conditionally independent, the likelihoods (5) for each replicate \mathbf{y}_k can be multiplied to yield the final likelihood

$$p(\mathbf{y}_1, \dots, \mathbf{y}_N; \boldsymbol{\beta}, \boldsymbol{\theta}_I, \boldsymbol{\theta}_Z, \mathbf{x}_1, \dots, \mathbf{x}_N) = \prod_{k=1}^N p(\mathbf{y}_k; \boldsymbol{\beta}_{0k}, \boldsymbol{\beta}_1, \boldsymbol{\theta}_I, \mathbf{x}_k), \quad (8)$$

where now $\boldsymbol{\beta}$ contains all the regression coefficients, i.e. $\boldsymbol{\beta} = (\beta_{01}, \dots, \beta_{0N}, \beta_1)$. To obtain an approximation of (8), the approximations (6) and (7) are applied to each pattern separately.

3.5 MCMC

Combining the likelihood (8) with the prior $p(\boldsymbol{\beta}, \boldsymbol{\theta}_I, \boldsymbol{\theta}_Z)$ yields the approximate posterior distribution. To sample from this distribution, we use Robust Adaptive Metropolis algorithm (Vihola 2012, 2020), which uses a Gaussian random-walk proposal distribution, whose covariance is updated adaptively. The limiting proposal covariance matches the shape of the posterior, such that an average acceptance rate of 0.234 is attained, following the theoretical findings presented e.g. in Roberts et al. (1997).

4 Simulation experiment

We made a simulation experiment to study the performance of the inference approach and the edge correction method suggested above. The point pattern \mathbf{x} was a realization of either a Poisson process or a regular Strauss process. The Strauss process (e.g. Illian et al. 2008) was included to see whether the edge correction based on the Poisson assumption of X would work even in a more regular case. We did not include any cluster process since in our application, the large tree patterns \mathbf{x} are regular. Also, based on a small simulation study (results not shown here), it is unlikely that the Poisson correction would work well when the \mathbf{x} pattern is strongly clustered. We did not include marks in the simulation experiment.

4.1 Set-up

The intensity parameters of the Poisson and Strauss processes were chosen such that they result in approximately 60 points in the observation window $W = [0, 40] \times [0, 40]$. In the Strauss process (parametrized as in Baddeley et al. 2015), the intensity related parameter was 0.06, the interaction strength 0.1, and the interaction radius 2, making the resulting patterns rather regular. The \mathbf{y} patterns were generated on W and the \mathbf{x} patterns on the extended window $W_{\text{ext}} = [-20, 60] \times [-20, 60]$ to be able to use plus sampling which represents the ideal situation where no imputation is needed as the complete pattern is known. The Gaussian kernel (1) was used as the influence kernel. Initially, the parameters of the competition field and of the Gaussian field were set to the estimates found in Sect. 5 and the intercept β_0 was chosen such that the resulting LGCP model would have 600 points on average. First we used the estimated values $\beta_1 = -0.7$ and $\theta = 2.1$, called “estimated” in Fig. 2. In addition, we used either the values $\beta_1 = -3$, and $\theta = 2.1$ corresponding to a much stronger effect of the influence kernel (β_1) (“strong” in Fig. 2) or the values $\beta_1 = -0.7$, and $\theta = 6$ corresponding to a much larger range of influence θ (“wide” in Fig. 2) than in the data. In all cases, $\sigma_Z = 1.6$ and $\rho_Z = 2.6$. We generated 100 replicates of each X process and one \mathbf{y} pattern for each \mathbf{x} . The random intensity of the Cox process was approximated by a piecewise constant function using $0.1 \text{ m} \times 0.1 \text{ m}$ cells.

We fitted the conditional LGCP model to the simulated point patterns. We discretized the observation windows to pixels of size $1 \text{ m} \times 1 \text{ m}$ and set weakly informative independent priors for all model parameters as follows: For the parameters in β , we used Gaussian distributions with mean zero and standard deviation 10. For the range parameters ρ_Z and θ , very small and very large values do not make sense based on the discretization and window used. Thus, we set the prior to be the Gamma distribution with shape parameter 2.4 and scale parameter 1.8, implying that approximately 90% of the prior probability is between 1 m and 10 m. Furthermore, the prior for the standard deviation of the Gaussian field σ_Z was the exponential distribution with expectation 10, slightly favoring small values.

For each point pattern, we then ran the MCMC scheme using (a) no edge correction, (b) the Poisson edge correction and (c) plus sampling edge correction with 100,000 updates using the true parameter values as the starting values. For each chain we

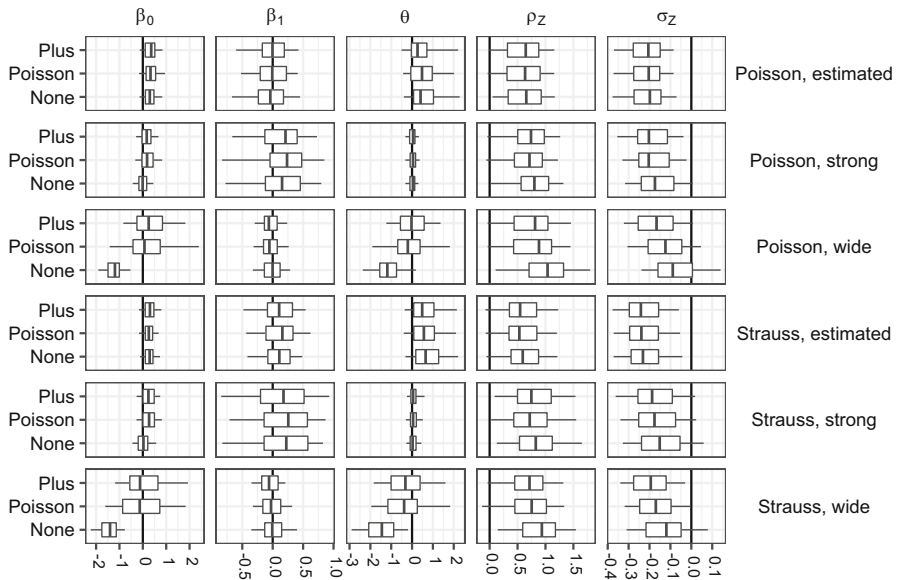


Fig. 2 Quantiles (0.05, 0.25, 0.5, 0.75, 0.95) of differences between posterior means and reference values. For each row of the figure, we display the X process and the competition effect on the right and within each subfigure, we label the three different edge corrections (left). The quantiles are based on 100 replicates

discarded 20,000 first samples as burn-in and saved every 10th sample. When influence was strong, most chains converged and mixed well. However, there were problems with mixing if the influence was not so strong. In this case the effective sample size was estimated to be less than 1000 in half of the chains. Upon closer inspection multi-modality was often the cause. We used posterior means of each chain in the comparisons. Using posterior modes led to identical conclusions.

4.2 Results

First, we investigated the performance of the Bayesian inference approach. To avoid edge effects, we estimated the parameters using plus sampling, utilizing the true pattern x in the extended window. Based on the distributions of the posterior means for the plus sampling method (see Plus in Fig. 2), we can see that the Bayesian MCMC approach with the approximations used performed reasonably well for the main parameters β_1 and θ . However, the less interesting random field parameters were clearly biased. As expected, the distribution of the X pattern did not affect the performance of the inference.

Second, we investigated the performance of the Poisson edge correction. An example of the expected intensity field with and without edge correction for the conditional LGCP model with the parameters estimated from the EVO02 pattern and Strauss pattern x is shown in top row of Fig. 3. It can be seen that the Poisson corrected and the plus sampling corrected intensities are quite similar to each other. The bottom row of Fig. 3 further shows the components of the influence field for the Poisson correction,

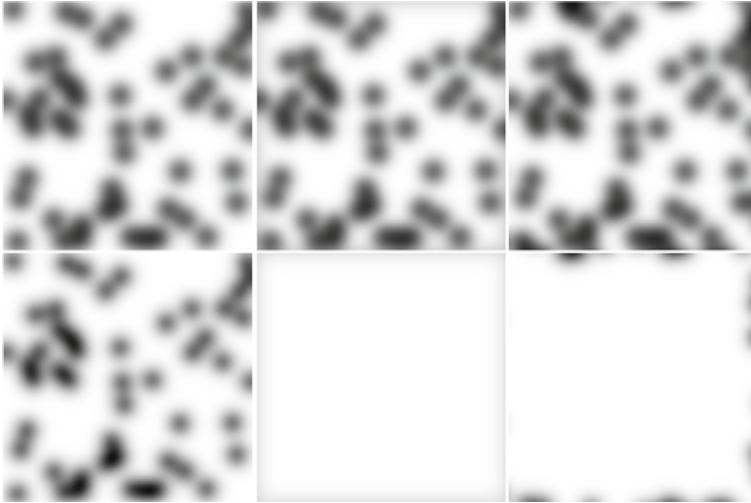


Fig. 3 Top row: expected intensity of the conditional LGCP with parameters estimated from the EVO02 pattern using no edge correction (left), Poisson correction (middle) and plus sampling (right). Bottom row: influence field induced by the observed points (left), expected influence field caused by the unobserved points under the Poisson assumption (middle) and influence field caused by the unobserved points (right). The x pattern is a realization of a Strauss process with interaction parameter 0.1, interaction range 2, and with on average 60 points. Dark color means low intensity/high influence

namely the contribution of the observed points (left) and the expected contribution of the unobserved points under the Poisson assumption (middle). The contribution of unobserved points is shown for comparison (right). The Poisson correction simply approximates the contribution of the unobserved points.

To assess the performance of the proposed edge correction method, we compared the posterior means of the model parameters β_0 , β_1 and θ , obtained by using plus sampling to the estimates obtained by using the Poisson correction and those obtained by using no edge correction. The distribution of the posterior means is shown in Fig. 2. It can be seen that the estimates of the different methods are very similar when the influence of the large trees was not too wide, for both X processes. However, when the influence was wide, the proposed Poisson correction produced estimates that were closer to the plus sampling based estimates than the uncorrected estimates were. The results were altogether very similar for the Poisson and Strauss processes. Thus, the edge correction plays a role if the range of influence of the x points on the intensity of Y is wide.

5 Application

The data shown in Fig. 1 have been collected on 40 m \times 40 m squares in southern Finland. They are part of a larger data set collected for studies on tree and stand development in managed, uneven-aged Norway spruce forests conducted under the ERIKA research project at the Natural Resources Institute Finland (Eerikäinen et al.

2007; Eerikäinen et al. 2014; Saksa and Valkonen 2011). Using the conditional LGCP model, we studied the effect of large trees x_i (black circles) to the seedling patterns y_i (red crosses). The patterns x_i consist of trees which had a vital crown with no damages and with a dbh at least 7 cm in 1991. Most trees (78% of trees, 70% basal area) were Norway spruces and the remaining ones either Scots pines or broadleaves. The seedlings were naturally generated with height at least 10 cm in 1996 and had reached this height after the data collection in 1991. The seedlings were mostly Norway spruces (98%).

We fitted the conditional LGCP model using different mark dependent and mark independent Gaussian influence fields: the full mark dependent model (2), the two reduced models where either of the mark specific parameters, namely δ or α were set to zero, the mark independent model (1), and a model without an influence field. The mark was always the dbh. We used the same discretization of the observation window (1 m \times 1 m pixel size) and the same priors as in the simulation experiment (Sect. 4.1). The pixel size 1 m \times 1 m was chosen because variations in smaller units are practically unimportant in forests. The priors for α and δ were both the exponential distribution with expectation 10. We then ran the MCMC scheme using the Poisson edge correction with 120,000 updates, leaving out the first 20,000 observations of the chains as the burn-in.

To compare the models, we used the posterior predictive model assessment based on various summary characteristics, namely the L -function (variance stabilizing version of Ripley's K), the empty space function F , and the nearest neighbour distribution function G summarizing the spatial pattern y and, to investigate the relationship between the large trees and seedlings, the cross L -function, L_{12} (e.g. Illian et al. 2008; Diggle 2013). We used the standard estimators of these functions with translational (L , L_{12}) and Kaplan-Meier edge correction methods (F , G) (Baddeley and Gill 1997). For each plot, we generated 10,000 patterns of seedlings from the posterior predictive distributions of the conditional LGCP models given the observed x and calculated the summary functions for the data and for each of the generated patterns. The posterior predictive simulations were made using a discretization with 0.2 m \times 0.2 m cell size.

Figure 4 shows the empirical L_{12} functions together with the 95% global extreme rank length envelopes (Myllymäki et al. 2017; Myllymäki and Mrkvička 2020) constructed from the L_{12} summary functions of the simulations of the fitted model with mark independent influence kernel (1) (shaded region), mark dependent influence kernel (2) (dotted lines), and no influence kernel (dashed line) separately for each plot. The observed L_{12} function is distinctly better covered by the envelopes based on the models with influence field than without. While the envelopes of the model without an influence field are centred around zero, i.e., no interaction between trees and seedlings, the empirical L_{12} functions have the tendency to go below zero in most plots, indicating repulsion or inhibition of trees and seedlings, and the envelopes of the models with influence kernels are shifted downwards as well. The difference between the two models with influence kernels is, however, minor. Other summary functions (L , F , G) produced very similar envelopes regardless of the type or lack of influence field, see figures in Appendix 1. The empirical functions were inside the envelopes, except the nearest neighbor distance distribution functions of four sample plots VES07, VES13,

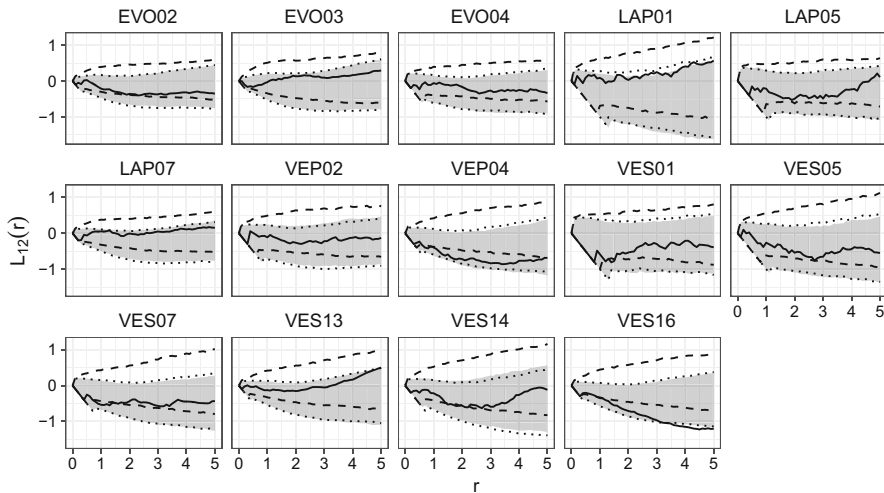


Fig. 4 Empirical L_{12} functions (solid line) together with the 95% global envelopes constructed from 10,000 simulations from the posterior predictive distribution of the fitted conditional LGCP models for the 14 plots in Fig. 1 with mark independent (1) (grey shade), mark dependent (2) (dotted lines), and no (dashed lines) influence

VES14 and VES16, which were slightly outside the envelopes at distances less than 1 m, i.e. less than the pixel size used in the discretization. This may suggest that the spatial distribution of the seedlings is not Poisson at a very small scale, but we did not investigate this further.

The envelopes for the models with mark dependent kernels with either δ or α set to zero are omitted because they were very similar to the envelopes of the other two influence kernels.

Based on the analysis above, it is clear that an influence kernel is needed. However, since all the models with an influence kernel fitted the data equally well, we report the results of the simplest model (1). The marginal posterior distributions of the model parameters of this model are shown in Fig. 6. The influence of the large trees on the seedlings (β_1) is clearly negative meaning that the seedlings avoid locations in the close vicinity of the large trees. The range of influence θ of the large trees was estimated to be around 2.1 m, indicating that the influence of a large tree decreases from its maximum influence (at the tree location) to 37% of it at distance 2.1 m from the tree, or to 5% of it at distance 3.6 m. However, there is a lot of unexplained variability, as the quite wide envelopes in Fig. 4 and Appendix 1 show.

Figure 5 shows for each plot one realization drawn from the posterior predictive distribution of the model with mark independent influence kernel. It is difficult to detect the relationship between trees and seedlings by eye, but one can compare the clustering of the seedling patterns to the observed patterns (Fig. 1). The patterns in Figs. 1 and 5 look rather similar, and according to the envelope tests (see Fig. 4 and Appendix 1) the model captures small scale structures up to 5 m distances.

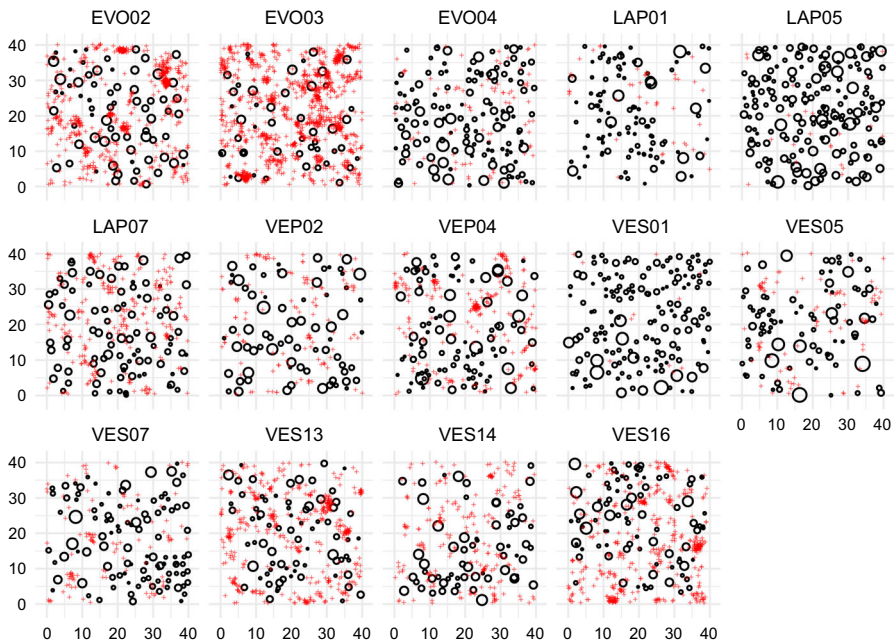
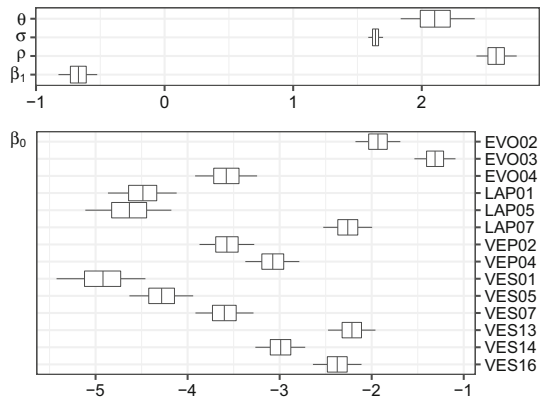


Fig. 5 Simulated seedling pattern (red crosses) and observed large trees (black circles, radius relative to dbh). The simulation was done using the posterior predictive distribution of the fitted conditional LGCP models for the 14 plots in Fig. 1 with mark independent influence of large trees

Fig. 6 Posterior quantiles (0.05, 0.25, 0.5, 0.75, 0.95) of the common parameters (top) and the sample plot specific intercepts β_0 (bottom)



6 Discussion

We proposed a LGCP model to investigate the effect of large trees on the intensity of seedlings under the presence of unexplained clustering. The influence of large trees was modeled by using parametric influence kernels around them. Our analysis suggests that tree regeneration is affected by the pattern of large trees in the studied data. Namely, the large trees were found to have negative effect on the seedling density in the vicinity of large trees. Further, the LGCP model could capture much of the unexplained clus-

tering. For parameter estimation, we constructed a Bayesian approach using MCMC and Laplace approximation. All computations were implemented in Julia language (Bezanson et al. 2017), while graphics were done using ggplot2 (Wickham 2016).

Estimation of the influence field parameters worked well in our simulation experiment and we did not observe any problems due to possible confounding between the influence field and the spatial random effect as reported in the literature (e.g. Dupont et al. 2020). However, the random field parameter estimates were biased. We suspect that this is caused by weak identifiability (cf. Anderes 2010; Zhang 2004) or discretization bias coupled with the Laplace approximation. We used replicates to help with the weak identifiability which was necessary for the plots with very few seedlings. In our further experiments with finer discretizations (results not shown), we observed issues with the approximation. In particular, when the number of points per cell was small, the approximate posterior appeared to degenerate. We are unaware of exact inference methods that would be feasible in our setting, but we are currently investigating new methods that could allow for more detailed investigation of this issue.

There are many alternative approaches to inference with log Gaussian Cox processes. For example, the R package INLA (Rue et al. 2009) uses Laplace approximation in a similar fashion as we did, but is somewhat restricted to linear models. Indeed, INLA can in principle accommodate our model using the `rgeneric` class (personal communication with Håvard Rue). However, we faced some computational difficulties in estimation. The R package `lqcp` (Taylor et al. 2015) uses MALA algorithm for efficient Bayesian inference for the full model including the latent field. The use of full MCMC might lead to better estimation of the random field parameters. However, the `lqcp` package is also restricted to linear models, whereby we were not able to apply it directly to our model. For Stan (Stan Development Team 2018) our random field model appears to be too complicated, however there are some recent advances see e.g. Margossian et al. (2020). Also, `inlabru` (Bachl et al. 2019) could be further investigated.

Since the large trees outside the sampling window may affect the intensity of the seedlings within the window, an edge correction assuming that the large trees were from a Poisson process was included in the estimation procedure. We demonstrated by a small simulation study that this edge correction can work well even when the large trees are from a regular process. Compared to no edge correction, it improved the parameter estimates when the range of influence was rather wide.

An obvious alternative strategy would be to include the locations of the unobserved trees outside the observation window to the MCMC estimation in a similar manner as considered in the inference for Neyman-Scott point processes (Møller and Waagepetersen 2004). This approach would allow incorporating prior information on the large tree process in the edge correction at the cost of increased complexity. Since we did not have important prior information and the effect of edge correction appeared minor, we did not explore this approach further. Ideas from Geyer (1999) or Gabriel et al. (2017) could be used to find further alternative edge correction methods.

Our proposed edge correction method can be efficiently implemented when the influence field is constructed as the sum of individual signals. In principle, a similar edge correction could be applied with different combination rules, such as product (e.g. Wu et al. 1985; Miina and Pukkala 2002) or max-fields (e.g. Penttinen and Niemi

2007). If also the influence kernel is binary, e.g. $c(h; \theta) = \mathbf{1}(h \leq \theta)$, then the max-field is split into two phases as well, namely influence and influence-free zones. However, we note that our proposed calculation of the expected influence of trees outside the observation window W does not generalize directly to other combination rules.

The models introduced in this paper could be useful even for natural (e.g. Abellanas and Pérez-Moreno 2018) or urban forests (Hauru et al. 2012). Furthermore, they could be used in an experimental setup, where realizations of seedlings would be generated for different large tree patterns and the success of regeneration evaluated by some spatial summary functions such as the empty space function. In a similar manner, the effect of different thinning strategies on the regeneration of trees could be evaluated.

It could be argued that, since the management was the same for all plots and the geographical differences minor, the plots should have had a common intercept parameter. However, this was clearly not the case due to large variation in numbers of seedlings from plot to plot. Since we used plot specific intercepts, it could be argued that all other parameters should be plot specific too. This was not possible in practice due to the problems with the random field parameters. We did not explore the alternative where the intercept and the influence field parameters would be plot specific but the random field parameters shared since the envelope tests already suggested adequate fit of the model.

The observed and simulated seedling patterns in the Figures 1 and 5, respectively, are very similar in several aspects while quite different in others. For example, the clusters seemed to be clustered in the VES13 plot. Although the envelope tests suggest that the model was able to capture the variability in the data, it depends on the specific application if the model is adequate. To best of our knowledge, this is the first point process model accounting for clustering of the seedlings in these uneven-aged forests.

There are many other factors than the vicinity of large trees that may affect the intensity of seedlings (Valkonen and Maguire 2005; Kuusinen et al. 2019). Therefore, the model could be further improved and unexplained variability decreased if some covariate information on local conditions within plots would be available to be included in the model. Further, plot level covariate effects could be added to the model in order to explain the numbers of seedlings in different plots. Finally, we modeled the influence of large trees as a function of the dbh, whose effect on the influence was, however, minor in our data. Other possibly useful marks could be the height, crown ratio or crown width of the tree, for example.

Acknowledgements The authors thank Helena Henttonen, Jari Hynynen and Sauli Valkonen (Luke) for discussions on the application, and Antti Penttinen for his comments on an earlier version of the manuscript. Further, the authors are grateful to Hilikka Ollikainen and Juhani Korhonen, who mastered the maintenance and measurements on the plots of the ERIKA data set. The authors wish to acknowledge CSC – IT Center for Science, Finland, for computational resources. We are grateful to the two anonymous reviewers for their constructive comments.

Funding Open access funding provided by Natural Resources Institute Finland (LUKE). MK, MM and MV were financially supported by the Academy of Finland (Project Numbers 306875, 327211, 295100 and 315619) and AS by the Swedish Research Council (VR 2018-03986).

Data availability The datasets analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Code availability The custom code is available from authors on request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix: Envelopes

See Appendix Figures 7, 8 and 9.

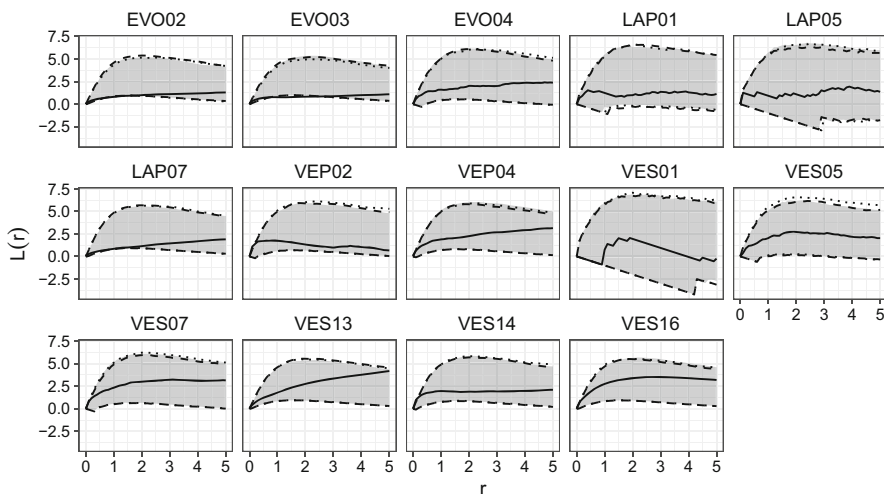


Fig. 7 Empirical L functions (solid line) together with the 95% global envelopes constructed from 10,000 simulations from the posterior predictive distribution of the fitted conditional LGCP models for the 14 plots in Fig. 1 with mark independent (grey shade), mark dependent (dashed lines), and no (dotted lines) influence

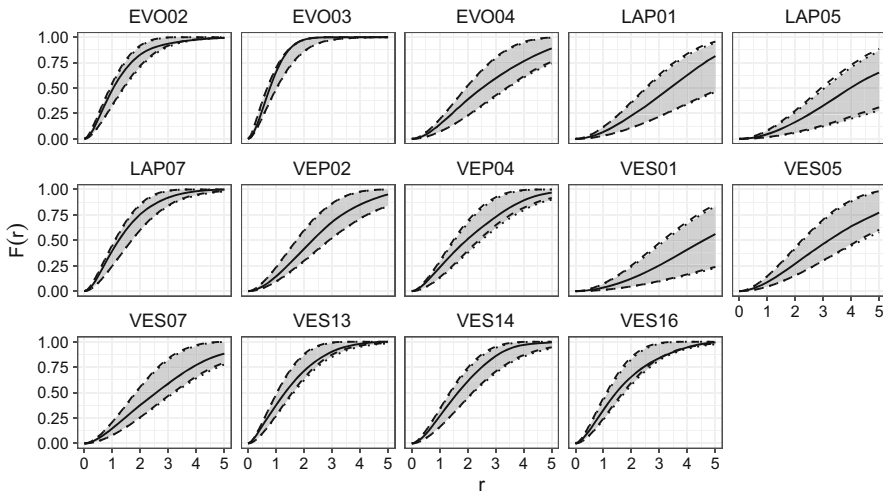


Fig. 8 Empirical empty space functions (solid line) together with the 95% global envelopes constructed from 10,000 simulations from the posterior predictive distribution of the fitted conditional LGCP models for the 14 plots in Fig. 1 with mark independent (grey shade), mark dependent (dashed lines), and no (dotted lines) influence

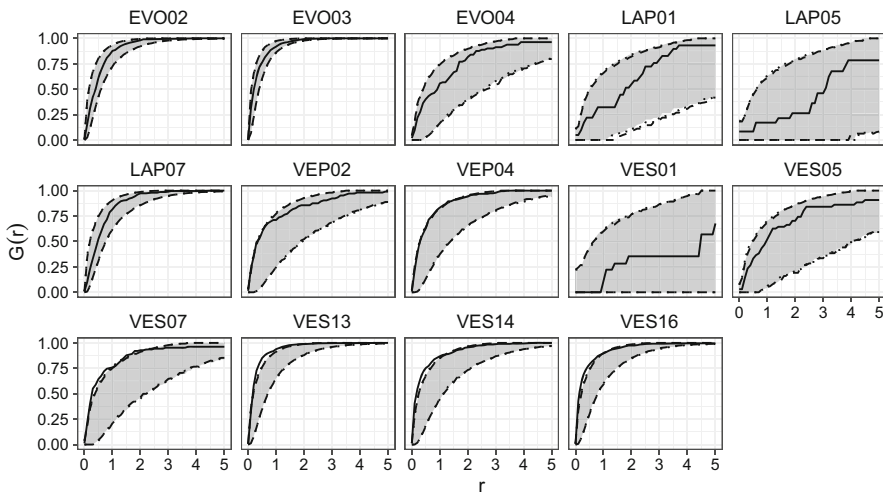


Fig. 9 Empirical nearest neighbor distance distribution functions (solid line) together with the 95% global envelopes constructed from 10,000 simulations from the posterior predictive distribution of the fitted conditional LGCP models for the 14 plots in Fig. 1 with mark independent (grey shade), mark dependent (dashed lines), and no (dotted lines) influence

References

- Abellanas B, Pérez-Moreno P (2018) Assessing spatial dynamics of a *Pinus nigra* subsp. *salzmannii* natural stand combining point and polygon patterns analysis. For Ecol Manag 424:136–153. <https://doi.org/10.1016/j.foreco.2018.04.050>
- Adler F (1996) A model of self-thinning through local competition. Proc Natl Acad Sci USA 93:9980–9984
- Anders E (2010) On the consistent separation of scale and variance for Gaussian random fields. Ann Stat 38(2):870–893. <https://doi.org/10.1214/09-AOS725>
- Bachl FE, Lindgren F, Borchers DL, Illian JB (2019) inlabru: an R package for Bayesian spatial modelling from ecological survey data. Methods Ecol Evol 10:760–766. <https://doi.org/10.1111/2041-210X.13168>
- Baddeley A, Gill RD (1997) Kaplan-Meier estimators of distance distributions for spatial point processes. Ann Stat 25(1):263–292
- Baddeley A, Rubak E, Turner R (2015) Spatial point patterns: methodology and applications with R. Chapman and Hall/CRC Press, London
- Banerjee S, Carlin BP, Gelfand AE (2004) Hierarchical modeling and analysis for spatial data, 1st edn. Chapman & Hall/CRC, Boca Raton
- Bezanson J, Edelman A, Karpinski S, Shah VB (2017) Julia: a fresh approach to numerical computing. SIAM Rev 59(1):65–98. <https://doi.org/10.1137/141000671>
- Chilés JP, Delfiner P (1999) Geostatistics: modeling spatial uncertainty. Wiley, New York
- Chiu SN, Stoyan D, Kendall WS, Mecke J (2013) Stochastic geometry and its applications, 3rd edn. Wiley, Chichester
- Cressie NAC (1993) Statistics for spatial data, revised edn. Wiley series in probability and mathematical statistics, Wiley, New York
- Dieckmann U, Law R, Metz J (2000) The geometry of ecological interactions: simplifying spatial complexity. Cambridge studies in adaptive dynamics. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511525537>
- Diggle PJ (2013) Statistical analysis of spatial and spatio-temporal point patterns, 3rd edn. CRC Press, Boca Raton
- Dupont E, Wood SN, Augustin N (2020) Spatial+: a novel approach to spatial confounding. [arXiv:2009.09420](https://arxiv.org/abs/2009.09420) [stat.ME]
- Eerikainen K, Valkonen S, Saksa T (2014) Ingrowth, survival and height growth of small trees in uneven-aged *picea abies* stands in southern Finland. For Ecosyst 1:5. <https://doi.org/10.1186/2197-5620-1-5>
- Eerikainen K, Miina J, Valkonen S (2007) Models for the regeneration establishment and the development of established seedlings in uneven-aged, Norway spruce dominated forest stands of southern Finland. For Ecol Manag 242(2):444–461. <https://doi.org/10.1016/j.foreco.2007.01.078>
- Frigo M, Johnson SG (2005) The design and implementation of FFTW3. Proc IEEE 93(2):216–231. (special issue on “Program Generation, Optimization, and Platform Adaptation”)
- Gabriel E, Coville J, Chadœuf J (2017) Estimating the intensity function of spatial point processes outside the observation window. Spat Stat 22:225–239. <https://doi.org/10.1016/j.spasta.2017.07.008>
- Genet A, Grabarnik P, Sekretenko O, Pothier D (2014) Incorporating the mechanisms underlying inter-tree competition into a random point process model to improve spatial tree pattern analysis in forestry. Ecol Model 288:143–154. <https://doi.org/10.1016/j.ecolmodel.2014.06.002>
- Geyer C (1999) Likelihood inference for spatial point processes: Likelihood and computation. In: Kendall W, Barndorff-Nielsen O, van Lieshout M (eds) Stochastic geometry. Chapman and Hall/CRC, Boca Raton, pp 141–172
- Grabarnik P, Särkkä A (2009) Modelling the spatial structure of forest stands by multivariate point processes with hierarchical interactions. Ecol Model 220(9–10):1232–1240. <https://doi.org/10.1016/j.ecolmodel.2009.02.021>
- Häbel H, Myllymäki M, Pommerening A (2019) New insights on the behaviour of alternative types of individual-based tree models for natural forests. Ecol Model 406:23–32. <https://doi.org/10.1016/j.ecolmodel.2019.02.013>
- Hauru K, Niemi A, Lehtvähirta S (2012) Spatial distribution of saplings in heavily worn urban forests: implications for regeneration and management. Urban For Urban Green 11(3):279–289. <https://doi.org/10.1016/j.ufug.2012.03.004>

- Högmander H, Särkkä A (1999) Multitype spatial point patterns with hierarchical interactions. *Biometrics* 55(4):1051–1058
- Illian J, Penttinen A, Stoyan H, Stoyan D (2008) Statistical analysis and modelling of spatial point patterns, 1st edn. Wiley, Chichester
- Illian JB, Møller J, Waagepetersen RP (2009) Hierarchical spatial point process analysis for a plant community with high biodiversity. *Environ Ecol Stat* 16(3):389–405. <https://doi.org/10.1007/s10651-007-0070-8>
- Kühlmann-Berenzon S, Heikkinen J, Särkkä A (2005) An additive edge correction for the influence potential of trees. *Biomet J* 47:517–526. <https://doi.org/10.1002/bimj.200310157>
- Kuuluvainen T, Pukkala T (1989) Effect of scots pine seed trees on the density of ground vegetation and tree seedlings. *Silva Fennica*. <https://doi.org/10.14214/sf.a15536>
- Kuusinen N, Valkonen S, Berninger F, Mäkelä A (2019) Seedling emergence in uneven-aged norway spruce stands in Finland. *Scand J Forest Res* 34(3):200–207. <https://doi.org/10.1080/02827581.2019.1575976>
- Lindgren F, Rue H, Lindström J (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J R Stat Soc* 73(4):423–498. <https://doi.org/10.1111/j.1467-9868.2011.00777.x>
- Margossian C, Vehtari A, Simpson D, Agrawal R (2020) Hamiltonian Monte Carlo using an adjoint-differentiated laplace approximation: Bayesian inference for latent Gaussian models and beyond. In: Thirty-fourth conference on neural information processing systems, Morgan Kaufmann Publishers, advances in neural information processing systems, conference on neural information processing systems, NeurIPS. Conference date: 06-12-2020 through 12-12-2020
- Miina J, Pukkala T (2002) Application of ecological field theory in distance-dependent growth modelling. *For Ecol Manag* 161(1):101–107. [https://doi.org/10.1016/S0378-1127\(01\)00489-3](https://doi.org/10.1016/S0378-1127(01)00489-3)
- Møller J, Waagepetersen RP (2004) Statistical inference and simulation for spatial point processes, 1st edn. Chapman & Hall/CRC, Boca Raton
- Møller J, Syversveen AR, Waagepetersen RP (1998) Log Gaussian Cox processes. *Scand J Stat* 25(3):451–482. <https://doi.org/10.1111/1467-9469.00115>
- Myllymäki M, Mrkvička T (2020) GET: Global envelopes in R. [arXiv:1911.06583](https://arxiv.org/abs/1911.06583) [stat.ME]
- Myllymäki M, Mrkvička T, Seijo H, Grabarnik P, Hahn U (2017) Global envelope tests for spatial processes. *J R Stat Soc* 79:381–404. <https://doi.org/10.1111/rssb.12172>
- Oppenheim AV, Schaffer RW, Buck JR (1999) Discrete-time signal processing, 2nd edn. Prentice-Hall Inc, USA
- Penttinen A, Niemi A (2007) On statistical inference for the random set generated cox process with set-marking. *Biomet J* 49(2):197–213. <https://doi.org/10.1002/bimj.200610272>
- Pommerening A, Grabarnik P (2019) Individual-based methods in forest ecology and management, 1st edn. Springer, New York. <https://doi.org/10.1007/978-3-030-24528-3>
- Pommerening A, Maleki K (2014) Differences between competition kernels and traditional size-ratio based competition indices used in forest ecology. *For Ecol Manag* 331:135–143. <https://doi.org/10.1016/j.foreco.2014.07.028>
- Pommerening A, Sánchez Meador AJ (2018) Tamm review: tree interactions between myth and reality. *For Ecol Manag* 424:164–176. <https://doi.org/10.1016/j.foreco.2018.04.051>
- Pommerening A, LeMay V, Stoyan D (2011) Model-based analysis of the influence of ecological processes on forest point pattern formation—a case study. *Ecol Model* 222(3):666–678. <https://doi.org/10.1016/j.ecolmodel.2010.10.019>
- Roberts GO, Gelman A, Gilks WR (1997) Weak convergence and optimal scaling of random walk metropolis algorithms. *Ann Appl Probab* 7(1):110–120. <https://doi.org/10.1214/aoap/1034625254>
- Rue H, Martino S, Chopin N (2009) Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *J R Stat Soc* 71:319–392
- Saksa T, Valkonen S (2011) Dynamics of seedling establishment and survival in uneven-aged boreal forests. *For Ecol Manag* 261(8):1409–1414. <https://doi.org/10.1016/j.foreco.2011.01.026>
- Schneider MK, Law R, Illian JB (2006) Quantification of neighbourhood-dependent plant growth by Bayesian hierarchical modelling. *J Ecol* 94(2):310–321. <https://doi.org/10.1111/j.1365-2745.2005.01079.x>
- Stan Development Team (2018) Stan modeling language users guide and reference manual, version 2.18.0. <http://mc-stan.org/>

- Taylor BM, Davies TM, Rowlingson BS, Diggle PJ (2015) Bayesian inference and data augmentation schemes for spatial, spatiotemporal and multivariate log-Gaussian Cox processes in R. *J Stat Softw* 63(7):1–48. <https://doi.org/10.18637/jss.v063.i07>
- Tierney L, Kadane JB (1986) Accurate approximations for posterior moments and marginal densities. *J Am Stat Assoc* 81(393):82–86. <https://doi.org/10.2307/2287970>
- Valkonen S, Maguire DA (2005) Relationship between seedbed properties and the emergence of spruce germinants in recently cut Norway spruce selection stands in southern Finland. *For Ecol Manag* 210(1):255–266. <https://doi.org/10.1016/j.foreco.2005.02.039>
- Vihola M (2012) Robust adaptive metropolis algorithm with coerced acceptance rate. *Stat Comput* 22:997–1008. <https://doi.org/10.1007/s11222-011-9269-5>
- Vihola M (2020) Ergonomic and reliable Bayesian inference with adaptive Markov chain Monte Carlo. In: Piegrorsch WW, Levine R, Zhang HH, Lee TCM (eds) *Handbook of computational statistics and data science*. Wiley, New York
- Wickham H (2016) *ggplot2: Elegant graphics for data analysis*. Springer-Verlag, New York. <https://ggplot2.tidyverse.org>
- Wu HI, Sharpe PJ, Walker J, Penridge LK (1985) Ecological field theory: a spatial analysis of resource interference among plants. *Ecol Model* 29(1):215–243. [https://doi.org/10.1016/0304-3800\(85\)90054-7](https://doi.org/10.1016/0304-3800(85)90054-7)
- Zhang H (2004) Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J Am Stat Assoc* 99(465):250–261. <https://doi.org/10.1198/016214504000000241>

Mikko Kuronen is a doctoral student. His field of research is spatial statistics and computational methods.

Aila Särkkä is Professor of Mathematical Statistics. She is working on spatial point process statistics, spatio-temporal modeling, and applications in forestry, materials science and neurology.

Matti Vihola is Associate Professor of Statistics. His research interests include computational statistics, in particular Monte Carlo methods for Bayesian inference.

Mari Myllymäki is Academy Research Fellow. Her research topics include spatial and spatio-temporal statistics and simulation-based methods with applications to forestry and life sciences.